

# Enhancement of the Prediction Accuracy of Pole Coordinates with Empirical Mode Decomposition

Zhao Danning<sup>1,2</sup>, Lei Yu<sup>1,3</sup>, Cai Hongbing<sup>1,3</sup>

(1. National Time Service Center, Chinese Academy of Sciences, Xi'an 710600, China;

2. University of Chinese Academy of Sciences, Beijing 100049, China;

3. Key Laboratory of Time and Frequency Primary Standards, Chinese Academy of Sciences, Xi'an 710600, China)

**Abstract:** This paper is aimed at separation treatment of low- and high-frequency components in polar motion forecasting and then improving time-series predictions. For the purpose, the empirical mode decomposition (EMD) is employed as a filter to extract low- and high-frequency signals from original pole coordinate data. The decomposition of the pole motion observations between 1986 and 2015 from the International Earth Rotation and Reference Systems Service (IERS) C04 series illustrates that the low-frequency fluctuations including inter-decadal, inter-annual, Chandler and annual wobbles and shorter-period high-frequency oscillations can be separated from the observed time-series by the EMD. On the basis of separation, the least-squares (LS) extrapolation of models for annual and Chandler wobbles and for the linear trend are used for deterministic prediction of the low-frequency fluctuations, while the autoregressive (AR) technology is applied to forecasting the high-frequency oscillations plus LS fitting residuals. Pole coordinate forecasts are calculated as the sum of LS extrapolation and AR predictions (LS+AR). We have evaluated the accuracy of our long-term predictions (up to 1 year in the future) in comparison with the IERS official predictions in terms of year-by-year statistics of 5 years. It is shown that the accuracy of the LS+AR method can be significantly improved using a combination of the EMD and LS+AR (EMD+LS+AR). Also, the proposed prediction strategy overall outperforms the IERS solutions. In addition, the predictions are compared with those from the Earth Orientation Parameters Prediction Comparison Campaign (EOP PCC). The comparison demonstrates that the developed scheme is a very accurate approach to predict polar motion. According to this study, it is concluded that polar motion predictions may be enhanced through separation treatment of different time-scale fluctuations and thus such processing seems to be necessary in pole coordinate prediction.

**Keywords:** Polar Motion; Prediction; Empirical Mode Decomposition; Least-squares (LS); Autoregressive (AR) Model

中图分类号: P227.1      文献标志码: A      文章编号:

## 1. Introduction

The Earth orientation parameters (EOP): universal time,  $UT - TAI$ , pole coordinates and nutation-precession corrections, describe the irregularities of the Earth's rotation. Technically, they are the parameters which provide the rotation of the International Terrestrial Reference System (ITRS) to the International Celestial Reference System (ICRS) as a function of time. EOP derived from modern space geodetic technologies, e.g., Very Long Baseline Interferometry (VLBI), Satellite Laser Ranging (SLR) and Global Navigation Satellite System (GNSS), are not available for users in real-time applications resulting from the complexity of data processing. Since rapid and accurate EOP predictions are required for various fields associated with reference frames such as interplanetary spacecraft tracking and navigation, positional

基金项目: 国家自然科学基金(11503031); 中国科学院西部之光项目(201491).

作者简介: 赵丹宁, 女, 博士生. 研究方向: 卫星导航信号监测评估与自主导航. Email: zhaodanning31@163.com

astronomy, geodesy and precise orbit determinations of artificial Earth satellites, it is essential to make high accuracy EOP predictions at least for a few days in the future<sup>[1]</sup>. The current paper focuses on predicting, pole coordinates up to 365 days in the future. Many different prediction approaches and techniques were used in the past to improve the prediction accuracy of polar motion. Most of these prediction approaches and techniques use only information within pole coordinate data, e.g., the combination of the least-squares (LS) extrapolation and autoregressive (AR) model (LS+AR)<sup>[2-3]</sup> and combination of the LS extrapolation and neural network (LS+NN)<sup>[4-5]</sup>. It has been reported that the LS+AR combination is one of the most accurate techniques for pole coordinate forecast<sup>[6]</sup>.

The basic problem in forecasting time-series is the necessity of separation treatment of high- and low-frequency variations<sup>[7]</sup>. This problem can be addressed by combination of the empirical mode decomposition (EMD) with the LS+AR model. Our work presents a hybrid EMD+LS+AR method for the enhancement of pole variation prediction. In this method, the low-frequency components consisting of the long-term trend, Chandler and annual wobbles determined by the EMD are first fitted and predicted by the LS extrapolation of the linear polynomial and harmonic model, and then the short-period and non-cyclic high-frequency components together with the LS fitting residuals are modelled by the AR technique. The subsequently predicted high-frequency components and LS residuals are added to the LS extrapolation in order to obtain the predicted values of pole coordinates. The comparison with the existing prediction methods and techniques shows that the presented EMD+LS+AR approach is a tremendous tool for pole coordinate prediction.

## 2. Prediction methodology

### 2.1 Empirical mode decomposition

The EMD can decompose a signal into a limited number of IMFs whose instantaneous amplitude and frequency are physically meaningful. It works on a simple assumption that, at any given time, the data may have many many coexisting simple oscillatory modes of remarkably different frequency, one superimposed on the other. Each component is defined as an IMF satisfying the following conditions<sup>[8]</sup>. (1) In whole data set, the number of zero crossings and number of extrema must either be equal or differ at most by one. (2) At any data point, the mean value of the envelope defined using the local minima and the envelope defined using the local maxima equals zero. With the above definition of an IMF, a signal can be decomposed into many

IMFs through the following sifting process.

Step 1: Identifying all the local extrema in time-series .

Step 2: Connecting all the local minima (maxima) by a cubic spline line to form the lower (upper) envelopes.

Step 3: Calculating the mean value of the lower and upper envelopes,  $\bar{e}_m$ , and then computing the difference between  $x$  and  $\bar{e}_m$  as a proto-IMF,  $h$ , i.e.,

$$(1)$$

Step 4: Examining whether  $h$  satisfies the above-mentioned conditions. If it does,  $h$  can be regarded as an IMF component. If not, the above sifting process is repeated. In the iterating process,  $x$  is treated as the data in the next iteration:

$$(2)$$

After  $k$  times of iterations,

$$(3)$$

and  $h$  becomes the first IMF, i.e.,

$$(4)$$

Step 5: Computing the residual and then treating it as the new data, i.e.,

$$(5)$$

Step 6: Repeatedly applying the decomposition procedure to all subsequent .

The procedure finally stops when the residual becomes a monotonic function or a function with only one extremum from which no more IMF can be extracted. Through the above procedure the original data are decomposed into  $N$  IMFs and one residual obtained,  $r$ , which can be either a constant or the adaptive trend, i.e.,

$$(6)$$

## 2.2 Least-squares extrapolation

In accordance with the EMD-based separated low-frequency components of pole variation data, a LS model may be fitted. As the sampling interval of the International Earth Rotation and Reference Systems Service (IERS) EOP C04 series are one solar day, we employ the following polynomial-sinusoidal function for LS fitting and extrapolation, i.e.,

$$(7)$$

where the bias and drift of the linear trend, amplitudes, and phases  $\phi_1$ ,  $\phi_2$ , of the annual and Chandler wobbles have to be estimated by the LS method using the separated low-frequency fluctuation data, and  $\phi_3$ .

The model is applied to deriving the LS fitting residuals as for all used to fit the LS model. In the process of prediction, is extrapolated to obtain a deterministic prediction of the low-frequency fluctuations.

### 2.3 Autoregressive model

Stationary time-series may be modelled by the AR model of order [AR()]. The technique allows one to capture the temporal dependence within stationary time-series. A stationary, zero-mean stochastic process (in our case ) is said to be an AR() if it satisfies

$$(8)$$

where is white noise with zero mean and variance , and are the AR coefficients.

The order can be selected by the information criterion (AIC)<sup>[9]</sup>. The AIC for an AR() is based on the following statistics:

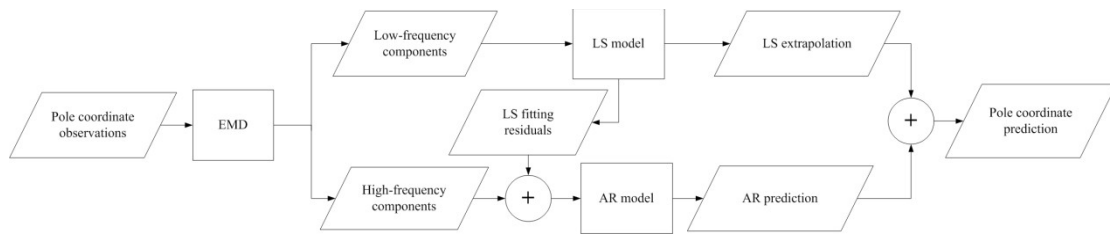
$$(9)$$

where is a likelihood function, is a vector of AR coefficients for an AR(), is an estimate of a variance and is a number of points of data. One selects which minimizes the AIC statistics. There are several approaches to estimate the AR coefficients; however, we use the Yule-Walker procedure.

The fitted AR model to is applied to forecasting the LS fitting residuals plus the high-frequency oscillations of polar motion using the linear prediction operator. The forecasts for several steps in the future are computed recursively.

### 2.4 Prediction strategy

For prediction of , pole coordinates, an integrated prediction strategy referred to as EMD+LS+AR is applied as described in the flowchart of Figure. 1. The presented strategy comprises the following stages. In the first stage, the low-frequency components including the annual wobble, Chandler wobble and long-term trend in pole coordinates and the high-frequency signals containing the shorter-period and non-cyclic fluctuations extracted respectively from original data using the EMD algorithm as stated in Section 2.2. In the second stage, the extracted low-frequency oscillations including the annual and Chandler wobbles are fitted and then extrapolated by the above LS model. In the next step, the AR technique are used to model the LS fitting residuals plus the separated high-frequency components of pole variations, i.e., . The subsequently forecasted values of are then added to the LS extrapolation in order to obtain the predictions of pole coordinates.



**Figure. 1** Flowchart of the EMD+LS+AR strategy for pole motion prediction

### 3. Numerical examples and results

#### 3.1 Data description

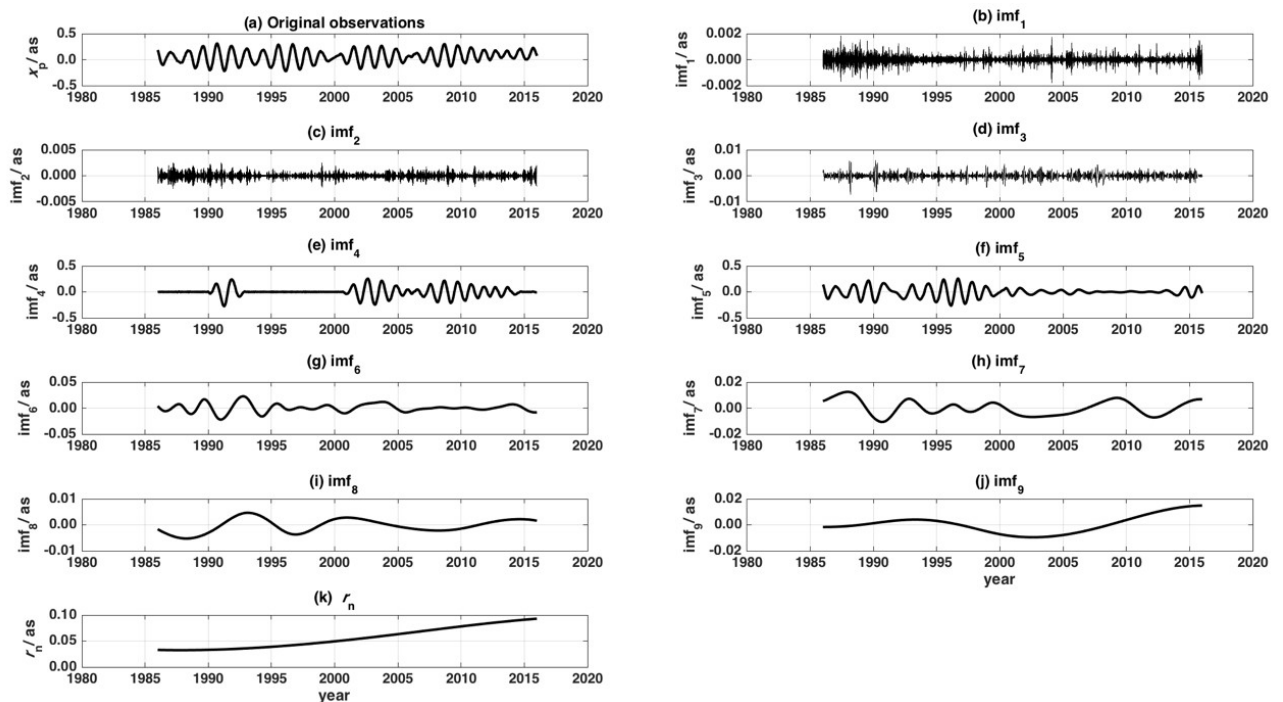
The IERS publishes daily values of EOP series with a latent of several days after the processing of observations of all advanced space geodetic techniques including GNSS, VLBI and SLR gathered at the permanent tracking stations distributed around the world. The EOP are estimated together with the station coordinates and velocities from each technique. The subsequently estimated EOP from all the above-mentioned techniques are then combined to derive the EOP C04 solutions<sup>[10]</sup>. These finally combined EOP solutions are regularly published at the official IERS website. In this work, the pole coordinate data from the IERS EOP 08 C04 series are used as data basis to carry out the numerical experiments. The present accuracy of pole coordinates is about  $30\mu\text{s}$  for the 08 C04 solutions.

#### 3.2 Separation of the low- and high-frequency oscillations in pole coordinates

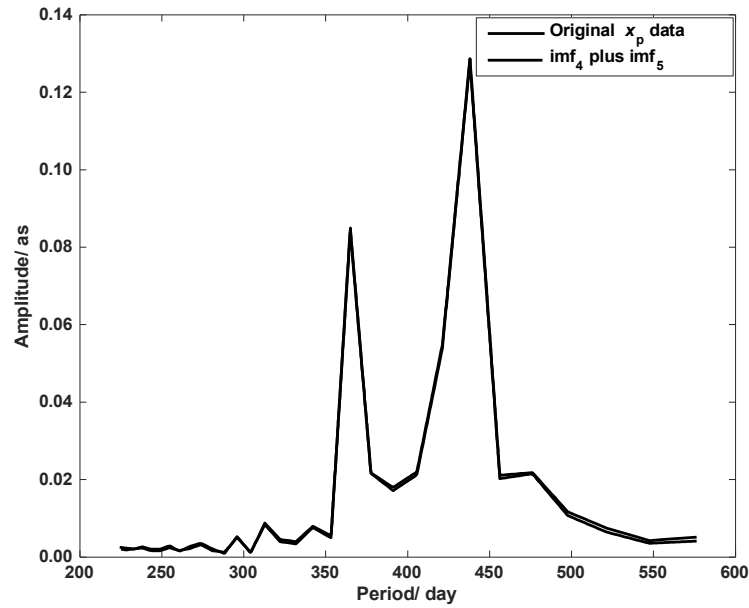
Taking the pole coordinate as an example, Figure. 2 shows the decomposed results of the 30-year pole variations from January 1, 1986 to December 31, 2015. It can be seen in Figure. 2 that the pole coordinate can be decomposed into a limited number of IMFs using the EMD algorithm, indicating that the oscillations of different time-scale in the component of polar motion are characterized by a small number of IMFs. As can be found in Figure.2, two of the modes found, namely and , show a noticeable seasonal oscillation in the range of a few hundreds of milliarcseconds, and thus can be considered as the two dominating oscillations along the and directions, namely annual wobble and Chandler wobble. In addition, we have found that the IMFs , , and with longer periods and the residuals arising from the EMD analysis can be identified as the inter-annual and inter-decadal signals and the long-term “natural” trend in the observations, and the modes , and with the smaller amplitudes represent the shorter-period oscillations and the very high-frequency non-wobbling components. It is noteworthy that due to the mode mixing problem in the EMD the mode separation may be difficult when the signal contains components having adjacent periods, as it

occurs for the Chandler and annual wobbles. Since here we only focus on separation of the low- and high-frequency oscillations rather than extraction of a particular component, however, this mode mixing problem dose not limit our analysis.

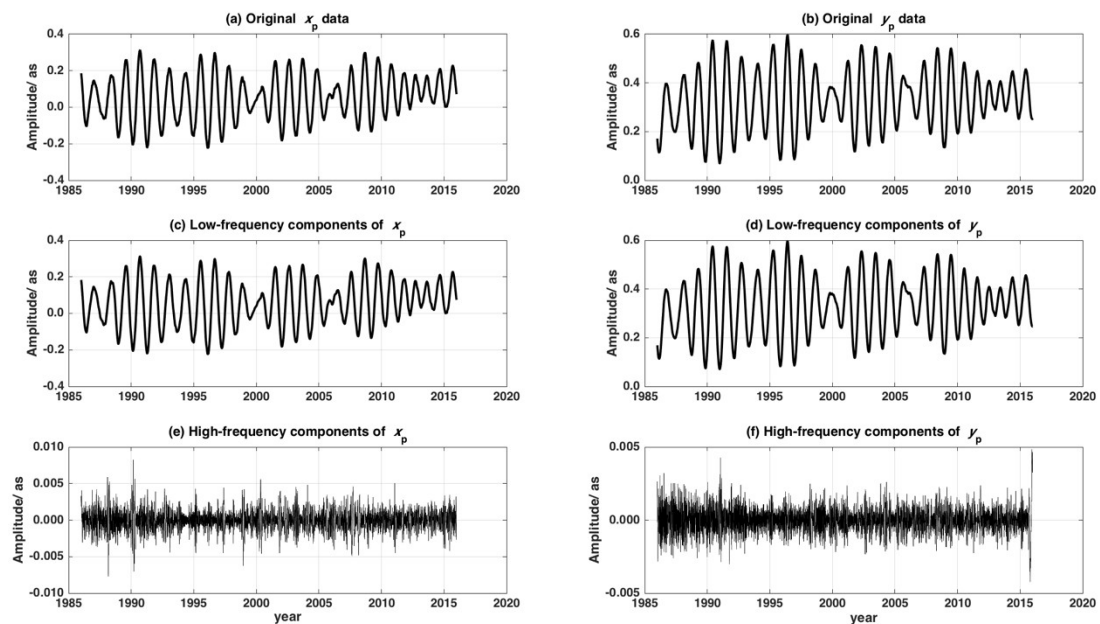
A spectral analysis of the IMFs plus based on fast Fourier transform (FFT) is exhibited in Figure. 3. In comparison, we also perform a similar analysis of the original data as given in Fig. 3. Visually, plus contains indeed contains the two dominant oscillating modes, i.e., Chandler and annual wobbles, with the same periods and amplitudes as the original observations do. Therefore, these finds confirm that the oscillating feature hidden in polar motion can be visually characterized by some oscillatory IMFs having particular physical meaning. To extract the high-frequency components of pole coordinates the IMFs , and are reconstructed, while the other modes are counted up to gain the low-frequency oscillations including the long-term trend, Chandler and annual wobbles. Figure.4 depicts the separated low-frequency time-series, high-frequency time-series of the , pole coordinates as well as the raw data. Their plots in Figure. 4 suggest that the low-frequency components including the Chandler wobble, annual wobble and long-term trend can be described by a low-order polynomial and harmonic model, while the AR technique is suitable for modelling the high-frequency terms.



**Figure.2**Original pole coordinate and its IMFs arising from the EMD algorithm



**Figure. 3** Spectral analysis of the oscillatory IMFs and original pole coordinates



**Figure. 4** Original , pole coordinates, the low- and high-frequency components

### 3.3 Prediction results

EOP 08 C04 series serve as data basis and the length of base data is 10 years. The predictions of , pole coordinates are generated by means of the two prediction schemes: LS+AR and EMD+LS+AR. The length of all predictions is equal to 365 days. In order to compare with the official predictions from the IERS Bulletin A, which contains EOP predictions for 1 year into the future at daily intervals and is issued by the IERS Rapid Service/Prediction Centre at the U.S. Naval Observatory



(USNO), all 365-day predictions of pole coordinate data are calculated every week at different starting prediction epochs from January 7, 2011 to December 31, 2015 because the IERS Bulletin A gives an advanced solution updated weekly. The comparison of the polar motion predictions is carried out by means of the analysis of: (1) the absolute values of differences between the pole coordinate data from the EOP 08 C04 series and their forecasts, and (2) mean absolute error (MAE) prediction errors defined by the following<sup>[6]</sup>

$$(10)$$

where  $t_0$  is the starting prediction epoch,  $n$  is the number of  $n$ -step predictions,  $y_{n,t_0}$  is the  $n$ -step prediction of polar motion and  $y_{t_0}$  is the corresponding observation of the IERS EOP C04 series. In our analysis,  $n=1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100$ .

The maximum absolute values of the differences between the pole coordinate data and their LS+AR predictions calculated at different starting epochs exceed 98 mas. The larger differences are for the selected polar motion forecasts with predictions lengths greater than 5 months (Figure. 5). In the case of the polar motion predictions published by the IERS Bulletin A, the maximum values of the differences between the data and IERS predictions do not exceed 91 mas and those differences are smaller than those computed for our LS+AR scheme (Figure. 6). For the tested data span, the combination of the EMD and LS+AR model produces polar motion predictions with maximum absolute differences of less than 87 mas and the extreme perturbations are significantly smaller than for the LS+AR and IERS solutions (Figure. 7).

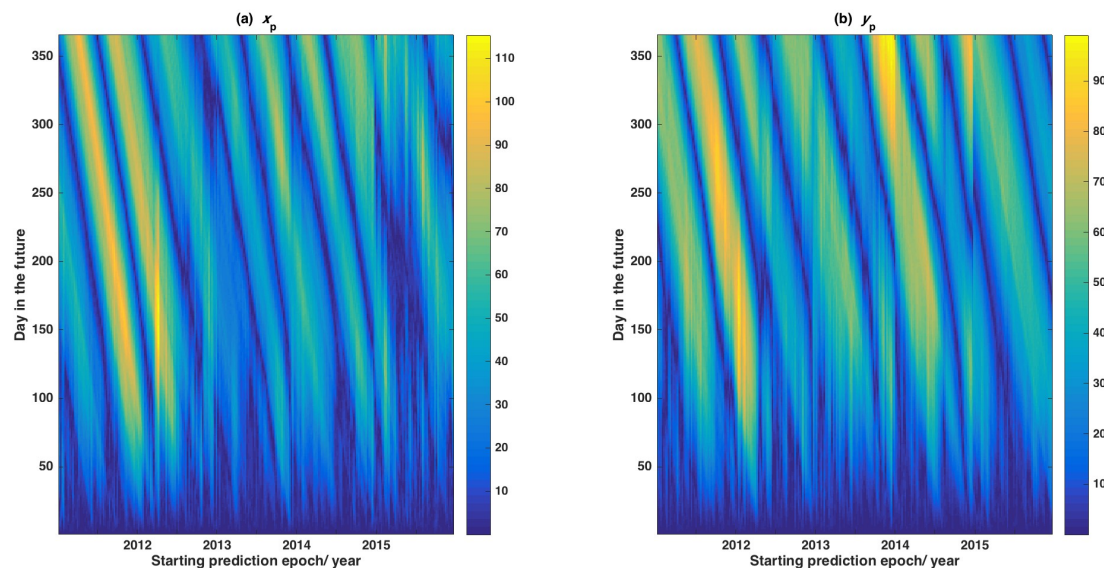
The representative statistic MAE as well as the corresponding error bars for our predictions in comparison with those for the IERS official predictions are shown year-by-year in Figure. 8. The results indicate that the accuracy of the LS+AR combination can be remarkably improved by incorporating the EMD into the prediction model, especially in long-term predictions. In contrast to the IERS predictions, our results obtained by the EMD+LS+AR approach appear to be comparable with those from the IERS for near-term prediction. However, for long-term prediction the proposed method noticeably outperforms the IERS solutions except individual prediction days, as can be seen apparently in Figure. 8.

In order to furthermore demonstrate the effectiveness of the developed EMD+LS+AR prediction strategy, the results of this strategy are also compared with

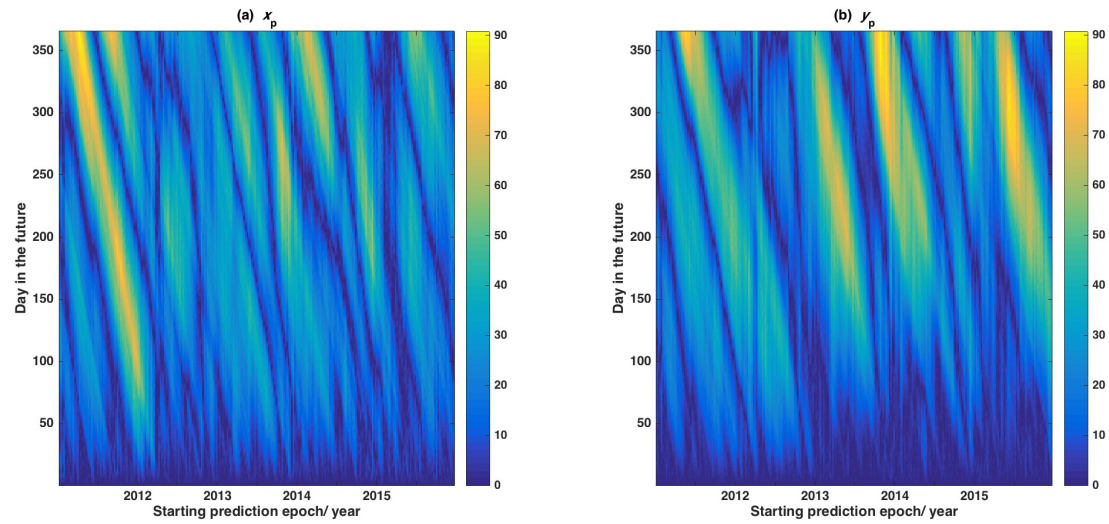


those from the EOP PCC lasting from October 1, 2005 to February 28, 2008, in which the prediction period and validation scheme had been clearly specified in advance. Under the same rules and conditions, the comparison with the contemporary prediction methods and techniques involved in the EOP PCC is shown in Figure. 9, in which the data time span is same, the statistics are referred to the EOP 05 C04 series and the MAE is used as the statistical measure. A list of participants and prediction methods can be found in [6].

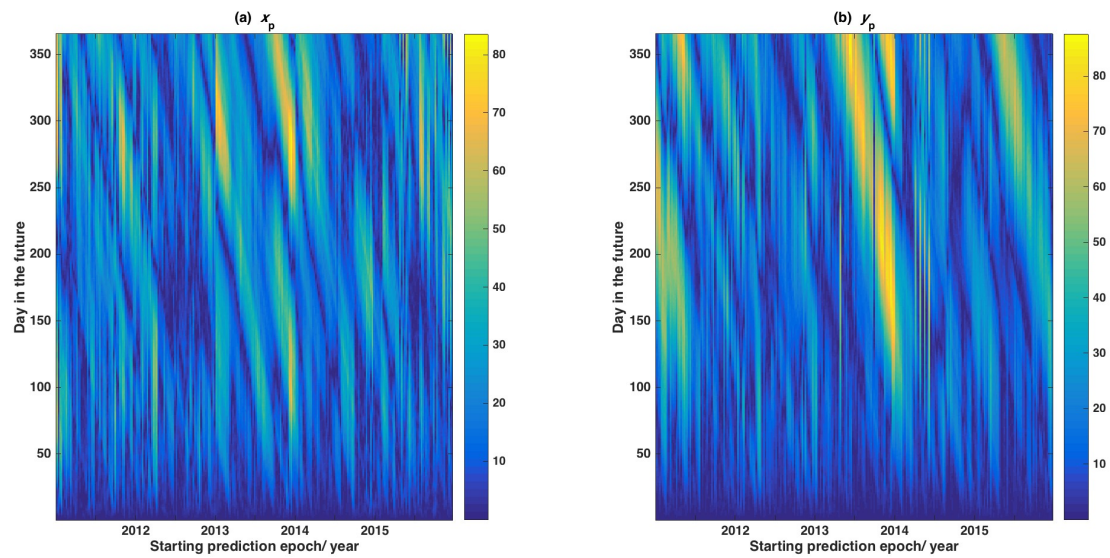
From the comparison it can be said that the error of the predictions for 1 ~ 10 days in the future by the EMD+LS+AR model is equal to the error of the most accurate prediction method for forecasting polar motion, namely the LS extrapolation of the harmonic model and AR prediction (Figure. 9(a)) and Figure. 9(b)). For short-term (up to 30 days in the future) prediction, the prediction accuracy of the developed method is comparable with or slightly worse than the accuracy of the three most accurate prediction techniques used (Figure. 9(c)) and Figure. 9(d)). As to long-term forecast ,(up to 1 year in the future) the predictions are inferior to the most accurate results, in particular those of data. However, our results are significantly more accurate than the predictions contributed by the other participants (Figure. 9(e)) and Figure. 9(f)).



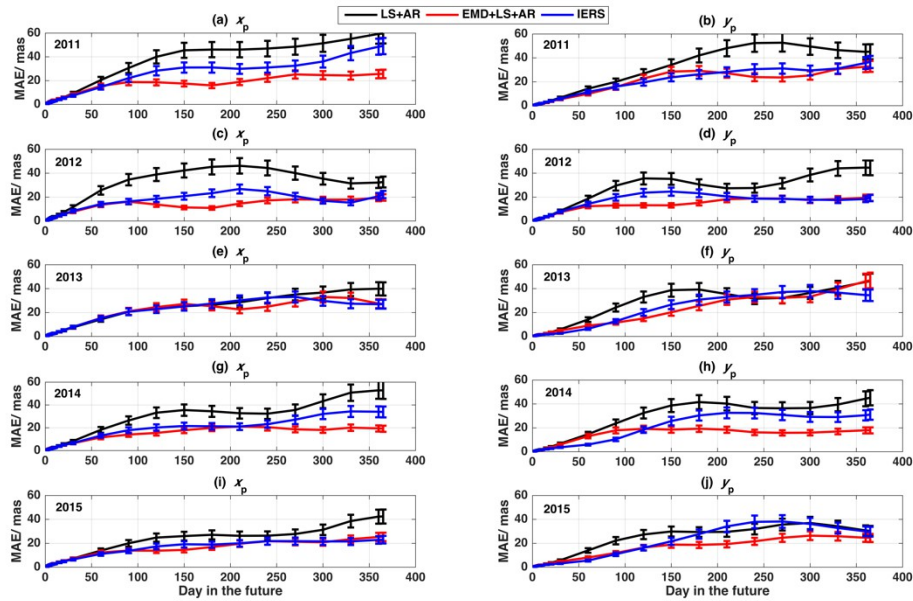
**Figure. 5** Absolute differences (in mas) between the pole coordinate data and their LS+AR predictions calculated at different starting epochs



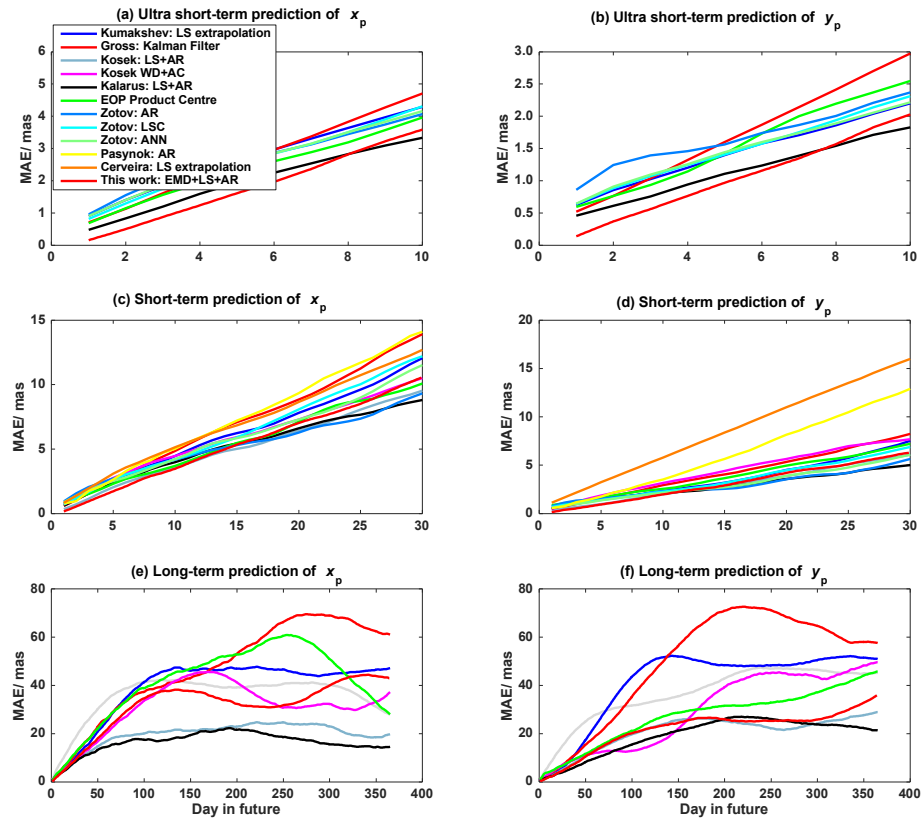
**Figure. 6** Absolute differences (in mas) between the pole coordinate data and their IERS predictions calculated at different starting epochs



**Figure. 7** Absolute differences (in mas) between the pole coordinate data and their EMD+LS+AR predictions calculated at different starting epochs



**Figure. 8** MAE for the LS+AR, EMD+LS+AR and IERS predictions of pole coordinates from 1 to 365 days in the future as well as the corresponding error bars



**Figure. 9** MAE for the EMD+LS+AR predictions and those contributed to the EOP PCC for different prediction days

#### 4. Summary and conclusion

Our investigations show that pole coordinate time-series were decomposed into a small number of IMFs representing different time-scale fluctuations, e.g., seasonal, inter-annual and inter-decadal fluctuations. Therefore, the EMD can be considered as a filter to extract specific oscillatory signals in polar motion. For the purpose of separation treatment of different components in time-series prediction, the EMD is used as a filtering method for separating the low- and high-frequency components of pole variations in our work. The results have indicated that both the shorter-period high-frequency variations and the low-frequency oscillations including the long-term trend, Chandler and annual wobbles can be extracted through reconstruction of the partial IMFs. Based on the findings, the LS extrapolation of the low-order polynomial and harmonic model and AR technology are recommended to predict the determined low-frequency and high-frequency components, respectively. The predictions of polar motion are generated by means of the combination of LS extrapolation and AR prediction. In order to evaluate the prediction accuracy of the proposed strategy, we have calculated the developed EMD+LS+AR as well as LS+AR predictions year-by-year at different starting prediction epochs between January 7, 2011 and December 31, 2015 in comparison with the IERS official forecasts. The figures of the MAE and absolute values of differences between pole coordinate data and their predictions illustrate that the presented hybrid approach indeed provides the more accurate predictions more than the conventional LS+AR combination, implying that the performance of the LS+AR method can be enhanced by involving the EMD into prediction model. Our results gained by the proposed method are better than the predictions released by the IERS as well, although those for individual prediction days are exception. In addition, the comparison between our results and the forecasts from the EOP PCC confirms that the EMD+LS+AR combination is a very powerful tool to predict pole coordinate data. The main conclusion is that the prediction error can be decreased when the EMD is incorporated into the LS+AR combination. Furthermore, we conclude that the separation treatment of different time-scale oscillations may enhance polar motion predictions and thus such processing seems to be of necessity in time-series prediction.

#### References

- [1] Gambis D, Luzum B. Earth rotation monitoring, UT1 determination and prediction [J]. Metrologia, 2011, 48(4): 65-170.

- [2] Xu X Q, Zhou Y H. EOP prediction using least square fitting and autoregressive filter over optimized data intervals [J]. Advances in Space Research, 2015, 56(10): 2248-2253.
- [3] Lei Yu, Cai Hongbing. Enhancing the prediction accuracy of the length of day change by eliminating the edge-effect of least squares fitting [J]. Astronomy Research and Technology, 2016, 13(4): 441-445.
- [4] Liao D C, Wang Q J, Zhou Y H, et al. Long-term prediction of the Earth orientation parameters by the artificial neural network technique[J]. Journal of Geodynamics, 2012, 62: 87-92.
- [5] Lei Yu, Cai Hongbing, Zhao Danning. Effects of training patterns on predictions of variations of length of day using an extreme learning machine neural network [J]. Astronomy Research and Technology, 2015, 12(3): 299-305.
- [6] Kalarus M, Schuh H, Kosek W, et al. Achievements of the Earth orientation parameters prediction comparison campaign [J]. Journal of Geodynamics, 2010, 84(10): 587-596.
- [7] Kosek W. Future improvements in EOP prediction[J]. Geodesy for Planet Earth, International Association of Geodesy Symposia, 2011, 136: 513-520.
- [8] Huang N E, Wu Z H. A review on Hilbert-Huang transform: method and its application to geophysical studies [J]. Reviews of Geophysics, 2018, 46(2): 1-23.
- [9] Akaike H. Autoregressive model fitting for control [J]. Annals of the Institute of Statistical Mathematics, 1971, 23: 163-180.
- [10] Bizouard C, Gambis D. The combined solution C04 for Earth orientation parameters consistent with international terrestrial reference frame 2005[J]. International Association of Geodesy Symposia, 2009, 134: 265-270.

## 利用经验模态分解提高极移预报精度

赵丹宁<sup>1,2</sup>, 雷雨<sup>1,3</sup>, 蔡宏兵<sup>1,3</sup>

(1.中国科学院国家授时中心, 西安 710600;

2.中国科学院大学, 北京 100049;

3. 中国科学院时间频率基准重点实验室, 西安 710600)

**摘要:** 经验模态分解 (Empirical Mode Decomposition, EMD) 是一种数据驱动的自适应非线性、非平稳信号分解方法。为提高极移预报精度, 将 EMD 应用于极移预报中。首先利用 EMD 方法对极移序列进行分解, 获得极移的高频分量和低频分量; 然后采用最小二乘 (Least Squares, LS) 外推模型对极移低频分量进行拟合, 获得 LS 拟合残差; 其次采用自回归 (Autoregressive, AR) 模型对极移高频分量和 LS 拟合残差之和进行建模预报; 最后将 LS 模型和 AR 模型外推值相加获得极移预报值。将 EMD 和 LS+AR 组合模型预报结果与 LS+AR 模型预报以及地球地球定向参数预报比较竞赛 (Earth Orientation Parameters Prediction Comparison Campaign, EOP PCC) 的预报结果进行比较,



结果表明，将 EMD 应用于极移预报中，可以明显改善极移预报精度。

关键词：极移; 预报; 经验模态分解; 最小二乘; 自回归模型